

WAVELET-BASED FACIAL MOTION CAPTURE FOR AVATAR ANIMATION

FIELD OF THE INVENTION

The present invention relates to dynamic facial feature sensing, and more particularly, to a vision-based motion capture system that allows real-time finding, tracking and classification of facial features for input into a graphics engine that animates an avatar.

BACKGROUND OF THE INVENTION

Virtual spaces filled with avatars are an attractive way to allow for the experience of a shared environment. However, existing shared environments generally lack facial feature sensing of sufficient quality to allow for the incarnation of a user, i.e., the endowment of an avatar with the likeness, expressions or gestures of the user. Quality facial feature sensing is a significant advantage because facial gestures are a primordial means of communications. Thus, the incarnation of a user augments the attractiveness of virtual spaces.

Existing methods of facial feature sensing typically use markers that are glued to a person's face. The use of markers for facial motion capture is cumbersome and has generally restricted the use of facial motion capture to high-cost applications such as movie production. Accordingly, there exists a significant need for a vision based motion capture systems that implements convenient and efficient facial feature sensing. The present invention satisfies this need.

SUMMARY OF THE INVENTION

The present invention is embodied in an apparatus, and related method, for sensing a person's facial movements, features or characteristic. The results of the facial sensing may be used to animate an avatar image. The avatar apparatus uses an image processing technique based on model graphs and bunch graphs that efficiently represent image features as jets composed of wavelet transforms at landmarks on a facial image corresponding to readily identifiable features. The sensing system allows tracking of a person's natural characteristics without any unnatural elements to interfere with the person's natural characteristics.

The feature sensing process operates on a sequence of image frames transforming each image frame using a wavelet transformation to generate a transformed image frame. Node locations associated with wavelets jets of a model graph to the transformed image frame are initialized by moving the model graph across the transformed image frame and placing the model graph at a location in the transformed image frame of maximum jet similarity between the wavelet jets at the node locations and the transformed image frame. The location of one or more node locations of the model graph is tracked between image frames. A tracked node is reinitialized if the node's position deviates beyond a predetermined position constraint between image frames.

In one embodiment of the invention, the facial feature finding may be based on elastic bunch graph matching for individualizing a head model. Also, the model graph for facial image analysis may include a plurality of location nodes (e.g., 18) associated with distinguishing features on a human face.

Other features and advantages of the present invention should be apparent from the following description of the preferred embodiments, taken in conjunction with the accompanying drawings, which illustrate, by way of example, the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of an avatar animation system and process, according to the invention.

FIG. 2 is block diagram of a facial feature sensing apparatus and process, according to the invention, for the avatar animation system and process of FIG. 1.

FIG. 3 is a block diagram of a video image processor for implementing the facial feature sensing apparatus of FIG. 2.

FIG. 4 is a flow diagram, with accompanying photographs, for illustrating a landmark finding technique of the facial feature sensing apparatus and system of FIG. 2.

FIG. 5 is a series of images showing processing of a facial image using Gabor wavelets, according to the invention.

FIG. 6 is a series of graphs showing the construction of a jet, image graph, and bunch graph using the wavelet processing technique of FIG. 5, according to the invention.

FIG. 7 is a diagram of a model graph, according to the invention, for processing facial images.

FIG. 8 includes two diagrams showing the use of wavelet processing to locate facial feature.

FIG. 9 is a flow diagram showing a tracking technique for tracking landmarks found by the landmark finding technique of FIG. 4.

FIG. 10 is a diagram of a Gaussian image pyramid technique for illustrating landmark tracking in one dimension.

FIG. 11 is a series of two facial images, with accompanying graphs of pose angle versus frame number, showing tracking of facial features over a sequence of 50 image frames.

FIG. 12 is a flow diagram, with accompanying photographs, for illustrating a pose estimation technique of the facial feature sensing apparatus and system of FIG. 2.

FIG. 13 is a schematic diagram of a face with extracted eye and mouth regions, for illustrating a course-to-fine landmark finding technique.

FIG. 14 are photographs showing the extraction of profile and facial features using the elastic bunch graph technique of FIG. 6.

FIG. 15 is a flow diagram showing the generation of a tagged personalized bunch graph along with a corresponding gallery of image patches that encompasses a variety of a person's expressions for avatar animation, according with the invention.

FIG. 16 is a flow diagram showing a technique for animating an avatar using image patches that are transmitted to a remote site, and that are selected at the remote site based on transmitted tags based on facial sensing associated with a person's current facial expressions.

FIG. 17 is a flow diagram showing rendering of a three-dimensional head image generated, based on facial feature positions and tags, using volume morphing integrated with dynamic texture generation.

FIG. 18 is a block diagram of an avatar animation system, according to the invention, that includes audio analysis for animating an avatar.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is embodied in an apparatus, and related method, for sensing a person's facial movements, features and characteristics and the like to generate and animate an avatar image based on the facial sensing. The